United Nations E/CN.3/2018/8

Economic and Social Council

Distr.: General
22 December 2017

Original: English

Statistical Commission
Forty-ninth session
6–9 March 2018
Item 3 (e) of the provisional agenda*
Items for discussion and decision: big data for official statistics

Report of the Global Working Group on Big Data for Official Statistics

Note by the Secretary-General

In accordance with Economic and Social Council decision 2017/228 and past practices, the Secretary-General has the honour to transmit the report of the Global Working Group on Big Data for Official Statistics. Over the last four years, the Working Group has identified many initiatives that aim to make new data sources, corresponding services and innovative applications accessible, and tried to harness their use in research and capacity-building for statistical production processes. The report responds to the request made by the Commission in its decision 48/105 and describes the progress made in the articulation and development of a global platform. This platform will be developed by the Working Group for global collaboration on trusted data, trusted methods, trusted partners and trusted learning, and is envisaged as a marketplace for sharing and developing trusted data, services and applications. The Working Group platform will operate as part of a federated network of autonomous platforms at the national, regional and global levels with defined interfaces to ensure interoperability and information-sharing. The Working Group task teams will collaborate on their use cases through the Working Group platform, which will be developed and operated under the auspices of the Statistical Commission. In the points for discussion contained in the last section of the report, the Commission is invited to comment on the articulation and development of the Working Group platform, on the progress made in 2017 and on the actions proposed for 2018.

* E/CN.3/2018/1.

17-22887 (E) 190118

Please recycle

I. Introduction

1. The Statistical Commission created the Global Working Group on Big Data for Official Statistics at its forty-fifth session, held in 2014. In accordance with its terms of reference (see E/CN.3/2015/4) and decision 46/101 of the Commission (see E/2015/24-E/CN.3/2015/40), the Working Group provides strategic vision, direction and the coordination of a global programme on big data for official statistics, including for indicators in the 2030 Agenda for Sustainable Development.

2. In its decision 48/105 (see E/2017/24-E/CN.3/2017/35), the Commission requested the Working Group to develop the business case and proof of concept for a global platform for collaboration on data, services and applications by applying an incremental and step-by-step approach.

3. In the present report, the Working Group responds to the above-mentioned request of the Commission and describes the progress made in the development of the business case and proof of concept for a Working Group platform of trusted data, trusted methods, trusted partners and trusted learning supported by core services, which will enable and advance the work of the Working Group task teams. Moreover, the Working Group reports on its annual management meeting, the fourth Global Conference on Big Data for Official Statistics, held in Bogota from 8 to 10 November 2017. The meeting was attended by senior managers from 20 national statistical agencies and 10 international organizations, complemented by senior representatives of several non-governmental organizations and private sector companies.[1] Its agenda covered the progress reports of the task teams, the proof of concept for the Working Group platform and the business case for the platform. The meeting agreed on the Bogota Declaration (see annex II to the present report).

4. There is a need for general platform facilities for the larger international community involved in big data for official statistics, as well as specific platform facilities for the Working Group itself, in particular its task teams. Therefore, it is expedient to make a distinction between the global network of platforms at the national, regional and global levels serving the larger international community, on the one hand, and the Working Group platform, defined as the platform of and for the Working Group itself, on the other. The Working Group platform is part of the global network and can be seen as its precursor.

5. The report highlights the main elements of the Bogota Declaration that describe the progressive articulation of the business case for the Working Group platform based on ongoing practices of the task teams. The Bogota Declaration also provides strategic direction and vision for the efforts of the Working Group to advance innovation and modernization of national statistical systems through trusted global data collaboratives developed by the task teams under the Working Group on the basis of trusted data, trusted methods, trusted partners and trusted learning.

6. The report describes the Bogota Declaration and the progress made in the articulation of the business case and proof of concept for the Working Group platform in sections II and III. Section IV shows the progress made by the various task teams of the Working Group and elaborates on the capacity-building prospects in this

[1] The Working Group meeting was attended by a large majority of its current members and by five new members, namely Switzerland, Germany, the International Monetary Fund, the African Development Bank and the Economic Commission for Africa. Saudi Arabia was also welcomed as a new member, even though it was unable to be represented at the meeting. The full updated list of Working Group members is available in annex I to the present report.

domain. Section V presents the outcome of the fourth Global Conference on Big Data for Official Statistics, while section VI describes the future actions to be taken by the Working Group to advance its work programme. The report concludes with points for consideration by the Statistical Commission.

II. The Bogota Declaration

7. The Bogota Declaration builds on the recommendations contained in the report entitled "A world that counts: mobilizing the data revolution for sustainable development", issued by the Secretary-General's Independent Expert Advisory Group on a Data Revolution for Sustainable Development, concerning: (a) technology, innovation and analysis to establish a network of data innovation networks for leveraging and sharing data and data research; (b) capacity-building and resources related to capacity-building, technology transfer, data literacy and resource mobilization through innovative financing mechanisms in partnership with the private sector; and (c) governance and leadership related to partnerships and coordination between governments, the private sector, non-governmental organizations, the media and academia to promote good practices and principles in data sharing, open data and data rights. The declaration also builds on the recommendations contained in the Cape Town Global Action Plan for Sustainable Development Data with respect to applying statistical standards and new data architecture for data sharing, exchange and integration, and to facilitating the use of new technology and new data sources in statistical production processes.

8. In the Bogota Declaration, the Working Group proposes to provide a major impetus for the strategic area of the Cape Town Global Action Plan on innovation and modernization of national statistical systems through trusted global collaboration and a global network of platforms, which is envisaged as a marketplace for sharing and developing core catalogues of services, data, metadata, methods, application programming interfaces, information technology tools and training materials. The global network operates as a federated network of platforms at the national, regional and global levels, which ensures interoperability and information-sharing among the platforms in the network through agreed and defined interfaces. The Working Group platform is intended for use by the Working Group task teams and is envisaged initially to meet the requirements for research and development in the use of multisource data. However, when relevant and legally possible, it will gradually incorporate the use of multisource data for statistical production, including the production of Sustainable Development Goal indicators. The Working Group platform is expected to be developed and operated under auspices of the Statistical Commission.

9. Also in the Bogota Declaration, the Working Group recommends fostering global collaboration, facilitated by the global network of platforms. The Working Group task teams can collaborate on their own platform within the federated network and should allow trusted data, trusted methods, trusted services and trusted applications to be shared as a public good, where useful and legally possible. Transparent partnership agreements will need to be developed with private- and public-sector organizations so that Working Group platform partners contribute and derive value through a business model which is individually sustainable for all stakeholders and ensures access to a trusted data and technology infrastructure.

10. The members of the Working Group agreed unanimously on the Bogota Declaration in person on 7 November and through subsequent exchange of emails. The full text of the declaration can be found in annex II.

III. The Working Group platform

11. The Bogota Declaration sets forth the details and drivers of the Working Group platform as a facilitator for sharing, exchanging and developing trusted data and metadata, trusted methods, trusted partners and trusted learning. This collaboration will gradually develop catalogues and libraries for data, metadata, methods, partners and learning on the Working Group platform (see figure).

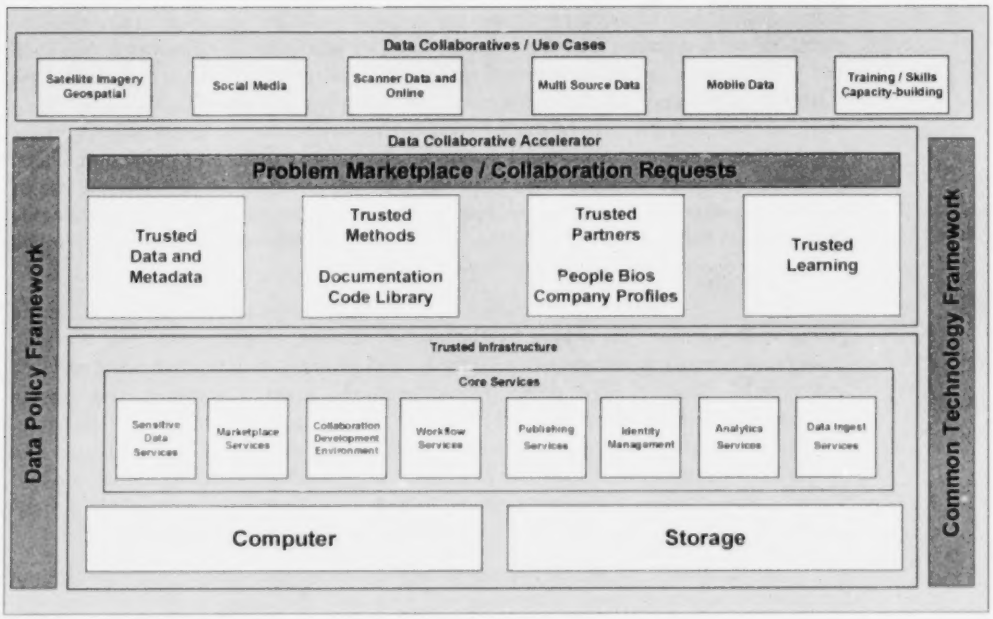
12. The Working Group platform will be part of an interconnected and federated network of platforms, which is based on the best practices of private and public big data initiatives. The Working Group platform offers technology infrastructure and constitutes a network for data innovation to facilitate global collaboration of the official statistical community on:

(a) The exchange of ideas and methods for processing, analysing and visualizing big data between official statisticians, data scientists and domain experts from the public and private sectors, with a focus on research and development building towards modernized statistical production;

(b) The sharing and exchange of trusted metadata, methods, services and applications for continuous development and reuse, including sharing trusted data where useful and legally possible. The sharing of sensitive data between trusted partners will be agreed upon bilaterally, while widely applicable open data sources can be generally made available;

(c) The development, jointly among the official statistical community, private sector technology companies and other communities, of a trusted data architecture so that different types of data can be shared safely and securely.

Architecture of the Working Group platform



13. As shown in the figure, the data policy framework of the Working Group platform is meant for data governance and information management and consists of a set of rules and regulations under which the Working Group platform must operate in storing, processing and disseminating information.[2] Following these rules will reduce the risks associated with inconsistently managed information, which will, in turn, reduce the cost of handling such risks and allow for increased flexibility in a changing technology and information environment over time. The common technology framework is a policy and technology framework meant to deliver data, accessible services and open source applications and associated application programming interfaces as a trusted public good to which all Working Group platform public and private sector network partners subscribe and adhere through contractual agreements.

14. The task teams of the Working Group would be the first partners to progressively develop the data and technology architecture of the Working Group platform on the basis of their requirements and real user demands and business cases. Other proofs of concept will include the trade data lake, which will initially consist of publicly available data on international trade and transport, such as trade data from the United Nations Comtrade database, data from the detailed database of the United Parcel Service, aviation data from the International Civil Aviation Organization and vessel tracking data. It is envisioned that gradually more sensitive data will be added to the platform, such as customs administration data (for example, data generated by the Automated System for Customs Data), which could be processed by trusted users, under strict protocols, directly in the data lake, making use of the latest tools and applications in the cloud server environment.

15. The Working Group platform is further expected to support capacity-building by compiling a library of trusted training materials, methods and software applications and conducting workshops on modernization of official statistics, the use of alternative data sources (such as big data) and the application of new tools, services and analytical techniques.

16. The development and maintenance of the platform will be undertaken under the auspices and guidance of the Commission, in support of national statistical systems of developed and developing countries. The implementation of the platform within the federated network of platforms will strengthen modern trusted data usage and information infrastructure for the global community of official statistics.

17. In further developing the platform and the underlying standards, methods, models and applications, the Working Group will have to make critical decisions on both technology and scientific matters. The Working Group has limited expertise in these areas. Therefore, the Working Group considered establishing two advisory groups, one on technology matters and one on scientific matters. These advisory groups would guide the Working Group on decisions regarding matters of technological development and scientific quality. It is foreseen that these groups will have a core group of advisers, which can be temporarily supplemented with additional advisers if certain matters so require. In addition, advisory groups could be established on legal matters and on fundraising to expedite the work of the task teams, including capacity-building, testing and experimentation in pilot countries.

[2] Some standards are already available at https://github.com/UNGlobalPlatform/standards.

IV. Working Group task teams and capacity-building

18. The Working Group has active task teams on the use of satellite imagery data, mobile phone data, social media data and scanner data and a task team on training, skills and capacity-building. The progress of each of these task teams is reported in the present section. A more detailed overview of their work and their training workshops will be posted on the Working Group website.

A. Satellite imagery data

19. The task team on satellite data has completed its handbook, which contains information about sources of earth observation data, methodologies for producing crop and other statistics using satellite imagery data, outlines of the task team's pilot projects and general guidance for the statistical community exploring the use of earth observation data for statistical purposes. The handbook will be reviewed by the advisory group on scientific matters, when established, before release on the public website of the Working Group.

20. On behalf of the Working Group, Professor Kerrie Mengersen and a colleague from the Queensland University of Technology in Brisbane, Australia, a scientist from the Department of Industry, Innovation and Science of the Government of Australia and a senior statistician of Statistics Canada developed a training course on the use of earth observation data for agricultural crop statistics. The course was conducted on 6 and 7 November 2017 in Bogota, prior to the fourth Global Conference on Big Data for Official Statistics. The 20 participants in the course came from statistical institutes from around the world.

21. The workshop gave advice on selecting data sources, on preparing raw satellite images for analysis and on different methods of producing statistical outputs. It also provided case studies on use of earth observation data to monitor key Sustainable Development Goal indicators, demonstrations in the use of software package "R" and hands-on assistance in producing statistical outputs from satellite imagery data. The practical focus was on crop-type identification and classification on the basis of Landsat 7 and ground reference data. The workshop led to a better understanding of the strengths and limitations of satellite data, the management and manipulation of satellite data and the application of satellite data for crop statistics.

B. Mobile phone data

22. The task team on mobile phone data finalized the first full draft of its handbook on the use of mobile phone data for official statistics, which describes in detail applications, data sources and methods. The handbook also includes business models for partnerships between national statistical offices and mobile operators for access to mobile phone data and concludes with two country cases from France and Indonesia. The handbook will be reviewed by the advisory group on scientific matters before release on the public website of the Working Group.

23. Prior to the fourth Global Conference on Big Data for Official Statistics and on behalf of the Working Group, two data scientists from Positium, a private sector intermediary, and one statistical methodologist from Eurostat conducted a workshop on the use of mobile phone data for official statistics for about 25 participants from national and international statistical institutes. The workshop used materials based on

projects with mobile operators across the European Union, the Middle East and Indonesia, resulting in the analysis of mobile data from 11 countries.

24. The trainers elaborated on the technical details of data preparation, data processing, and data modelling using hands-on examples, such as applications to model population movement, migration patterns, mobility, displacement due to diseases, tourism and real-time crowd mapping. Data related to these areas can be collected very quickly, in some cases in real time, from the mobile phone operators. Access to the data varies from country to country based on different legal and political frameworks. The workshop concluded with a discussion of possible road maps for embedding the use of mobile phone data in the official statistics of the participating countries.

C. Social media data

25. The task team on social media data is also in the process of completing a handbook on the use of social media data for opinion and sentiment indicators. It includes descriptions of various social media data sources (such as Twitter and Facebook), explanations of methods and techniques to analyse such data sources and descriptions of applications. Once the handbook is finalized, it will be reviewed by the Working Group advisory group on scientific matters before release on the public website of the Working Group.

26. Prior to the fourth Global Conference on Big Data for Official Statistics and on behalf of the Working Group, statisticians from the National Institute of Statistics and Geography of Mexico (INEGI) and the National Administrative Department of Statistics of Colombia (DANE) and a data scientist from Statistics Netherlands conducted a workshop with materials based on projects covering mobility analyses, sentiment analyses and subjective poverty measurement using mostly Twitter and Facebook data. The trainers explained the process of obtaining social media and web scraping data through application programming interfaces, the software and programming languages used to clean and process the data and the latest methodologies for data analytics, including machine learning.

27. Understanding of the scope, nature and coverage of social media data is important in these analyses. Not all social media platforms are equally popular among different age groups. Analysts need to have an understanding of how many users there are and their socioeconomic profile, i.e. they need to understand the population of their study. The behaviour of the users (e.g. frequency of posting) and the reach of the posts (whether public or private) can also influence the opinions which are expressed.

D. Scanner data

28. A new task team on scanner data was established in April 2017. It hit the ground running by making available the software code and statistical methods developed by the Bureau of Statistics of Australia, Statistics New Zealand and Statistics Netherlands for the use of scanner data in the production of consumer price indices. These statistical methods and software code were sharable and well-documented. The methods have been reviewed and released as open-source applications using software package "R". The task team made an explicit point of keeping the Intersecretariat Working Group on Price Statistics informed about the progress of its work.

29. Members of this task team have also started drafting a guide to explain and demonstrate a broad set of the methods for using scanner data in the calculation of consumer price indices. The availability of the guide and open source software will allow other statistical offices to experiment and test scanner data for potential use in their statistical production process, along with web-scraped and survey data.

30. On behalf of the Working Group, statisticians from the Bureau of Statistics of Australia, Statistics Belgium, Statistics Canada and Statistics Denmark conducted a training workshop on scanner and the use of online data for the compilation of consumer price indices the evening prior to the fourth Global Conference on Big Data for Official Statistics. The workshop provided information on how to access online data, select data sources and prepare raw data for use and described the status of scanner data implementation in different countries. The workshop also included a live demonstration using the consumer price index software written with software package "R" and provided hands-on assistance in scraping online data from websites. This led to a better understanding of the strength and limitations of the use of scanner data and online data and the strengths and weaknesses of several index methodologies for the compilation of consumer price indices.

E. Training, skills and capacity-building

31. The task team on training, skills and capacity-building organized its work around developing methods and tools to assess the needs for big data skills in national statistical systems, coordinating Working Group training courses, developing a modular training curriculum and cooperating within the global network of training programmes on the use of big data for official statistics. The task team worked to enrol representatives from developing countries in the European Statistical Training Programme (ESTP) to ensure they were properly represented.

32. Experiencing high demand from European statistical offices for ESTP courses related to the use of big data for official statistics, the task team accelerated the development of its own training courses with the support of other Working Group task teams. The success of the four training workshops organized by the task teams in Bogota in November 2017 demonstrates that the Working Group is able to offer a trusted learning programme with the support of member countries of the task teams. It is expected that a trusted learning curriculum will be further developed and announced for 2018 in collaboration with members of the task teams and the Working Group management group.

V. Outcome of the fourth Global Conference on Big Data for Official Statistics

33. The fourth Global Conference on Big Data for Official Statistics was held at the Ministry of Information and Communications Technologies of Colombia in Bogota from 8 to 10 November 2017. The sessions were attended on average by about 280 statisticians, data scientists and information technology specialists from around the world, with 350 participants turning out for the first day of the Conference. The Conference was organized by the Working Group in close cooperation with DANE, the national statistical office of Colombia.

34. The overarching theme of the conference was the use of "trusted data collaboratives" to share, exchange and develop data, services and applications for the

modernization of national statistical systems. In his keynote speech,[3] Alexandre Barbosa, Director of the Regional Centre for Studies on the Development of the Information Society in Brazil, clearly laid out three strategic areas for the innovation and modernization of national statistical systems: (a) institutional reform with respect to multi-stakeholder partnerships and the leading role of the national statistical offices; (b) new standard setting for open data, data security and interoperability; and (c) the use of new big data sources and corresponding technologies, including data storage, data integration and machine learning and artificial intelligence. He recognized that the use of big data had human, technical and technological challenges. However, the major impediment was institutional in nature. He highlighted good practices in addressing those challenges through the establishment of multi-stakeholder "trusted data collaboratives" by, for example, national data centres and data science campuses for research and development and statistical production. At the international level, a platform of the Global Working Group on Big Data for Official Statistics would complement the emerging network of national and regional platforms for the innovation and modernization of official statistics.

35. The Conference approached its main theme starting from the need for innovation and modernization of the national statistical systems, forging a strong link with the Cape Town Global Action Plan regarding the monitoring of progress on the 2030 Agenda for Sustainable Development through the use of new data and innovative technology. New collaborative initiatives were presented for establishing national data centres and science campuses and creating city data centres in selected municipalities for the purposes of harmonization and value creation by the national statistical system. Subsequently, panel sessions focused on broadening participation in the existing global collaboratives led by Working Group tasks teams in the use of earth observation data, scanner data and mobile phone data. Each of the panels emphasized the role of the statistical community in setting standards for trusted data, methods, partners and learning in the use of the data sources. Means should also be established to enable access to and use of the standards through a Working Group platform. A dedicated panel session on standards, consisting of representatives of the private and public sector, confirmed the need for defined principles and standards on encryption, data security and data protection for secure information sharing, which are subject to continuous technological update.

36. In the closing session of the Conference, a proof of concept of the envisaged data and technology infrastructure of the Working Group platform was presented by the Office of National Statistics of the United Kingdom of Great Britain and Northern Ireland. The data and technology model is built on the concept of trusted data, trusted methods, trusted partners and trusted learning to connect and create value for users and stakeholders and is supported by large sets of agreed core services, which are shareable within the international collaboration. Specific examples were shown, supported by video demonstrations covering some of the platform's core supporting services relating to sensitive data services, software development, publishing services, analytics services and data ingestion services. All of the proposed core services are based on open source software and are therefore freely available.

37. The Conference concluded with the sharing of views on the Bogotá Declaration[4] and the expression of commitments to collaborate on the Working Group platform through the existing and new task teams.

[3] The text of the speech is available from https://unstats.un.org/unsd/bigdata/conferences/2017/.../default.asp.

VI. Proposed actions of the Working Group

38. The Working Group has created a committee dedicated to the articulation and development of the Working Group platform, and collaboration on the platform has been organized through the Working Group task teams acting as global data collaboratives. Task teams have a known reporting mechanism within the Working Group and new data collaboratives will be created as new task teams under the Working Group. Short-term actions for the Working Group are completion of proofs of concept for the Working Group platform, upgrading of the Working Group website to provide a better vehicle for communication and advancement of the Working Group's capacity-building programme.

A. Adapting the governance structure

39. The Working Group has agreed to have two co-chairs because of the increased workload, in general, and the specialized knowledge required in building the Working Group platform, in particular. It has further been agreed that Statistics Denmark and the Office of National Statistics of the United Kingdom will be the co-chairs of the Working Group. The Statistics Division of the Department of Economic and Social Affairs of the Secretariat, as secretariat, will keep the co-chairs informed about all activities of the Working Group.

40. Statistics Denmark will oversee the Working Group's programme of work relating to task teams. The current management team consists mainly of the task team leaders, led by Statistics Denmark. This team has been renamed the committee on global data collaboratives, given that the task teams are functioning as global data collaboratives, and will be chaired by Statistics Denmark. The existing committee on the global platform will continue its work under the leadership of the Office of National Statistics of the United Kingdom. This committee will gradually develop the business case, proof of concept and communications for the Working Group platform, with a focus on its data and technology framework within a federated network of platforms. On a regular basis the Statistics Division will inform the full Working Group membership about the developments led by the two Working Group committees.

41. Two advisory groups are being considered: an advisory group on scientific matters chaired by Statistics Denmark and an advisory group on technology matters chaired by the Office of National Statistics of the United Kingdom. Members of the advisory groups would be invited on a personal basis, determined by their accomplishments and experience in their respective fields. Appropriate terms of reference for the groups and the composition of members would need to be determined. The advisory group on scientific matters would, as one of its first tasks, review the existing handbooks prepared by the task teams on satellite imagery data, mobile phone data and social media data, while the advisory group on technology matters would advise on the set-up of the Working Group platform and the interoperability of the platforms in the federated network.

42. For the progressive detailing of the business case for the Working Group platform based on the experiences of the task teams, the task teams will be asked to elaborate on legal aspects of their data collaboration, such as access to and use of privately owned data, collaboration with the private sector, data integration and the like. Also, the task teams will be asked to elaborate on the financial implications of the operation and maintenance of their data collaboration, including capacity-

building, testing and experimentation in pilot countries. Depending on the progress made by the task teams, specific advisory groups may be established for legal and funding matters.

B. Broadening the tasks of existing task teams and establishing new ones

43. Existing task teams will be requested to broaden the scope of their work by explicitly (a) including the sharing of proprietary data and methods, (b) developing training workshops for participants in developing countries and (c) expanding their partnerships with the private sector. Moreover, tasks teams are expected to start using the Working Group platform and its core services to share information. In light of the new work programme, the tasks teams will be asked to review their terms of reference early in 2018 and report back to the Working Group committee on global data collaboratives.

44. New task teams will be created, notably a task team on the trade data lake led by the Statistics Division, a task team on multisource data led by Statistics Denmark and a task team on data security and encryption led by the Office of National Statistics of the United Kingdom. The Working Group will continue to explore the creation of other teams on thematic issues such as economic and financial statistics, the digital economy and climate change, working with trusted partners. A critical aspect of establishing the new task teams will be to determine, for each, the chair and committed team members who will make their participation part of the regular programme of work within their organizations and who are willing to work using agile development processes with a dedicated approach.

C. Advancing the trusted Working Group platform

45. In 2018–2019 the Working Group committee on the global platform will work, together with the advisory group on technology, to develop a prototype platform whose use will be piloted by the existing Working Group task teams for initial testing. The Working Group platform will be built on a flexible cloud-based infrastructure initially to allow public data, services and applications to be shared as a public good where useful and legally possible.

46. As soon as is feasible, the Working Group platform will be opened up to trusted partners to facilitate the development of trusted data collaboratives. The initial principles and processes for identifying trusted partners will be determined by the committee on the global platform.

47. On the basis of agile development principles, the Working Group platform will be iterated to develop and deliver a marketplace which will offer trusted partners the ability to (a) collaborate among themselves to innovate and explore new data sources, (b) test and share different code, statistical and data science methods, (c) participate in learning and capability-building activities of benefit to their own organizations, (d) choose among a variety of core services and compute and storage solutions (including federating their own compute and storage facilities as part of the platform), and (e) collaborate to agree which data sources, code and methods are considered "trusted", i.e. tested and agreed robust enough for general-purpose use in the delivery of statistics and indicators.

48. Working with the advisory group on technology, the committee on the global platform will seek to develop the necessary technical solutions to extend the Working

Group platform to allow differentiated access to data of different types and classes, by user and use-case, by 2019–2020. This will enable data providers to share more sensitive data, not just public data, in such a way that they can be confident about which users have access to their data and for what purposes. This proof of concept of differentiated access to data and data sharing will be undertaken in close consultation with countries, taking into account their legislative frameworks to allow them to participate in the global network of platforms at a level which suits their local situation.

D. Developing the work programme on trusted learning

49. The four training workshops organized in Bogota demonstrated that the Working Group is able to offer a course curriculum on the use of new data sources for the compilation of official statistics. The courses were conducted with strong support from the members of the Working Group task teams. It is expected that this capacity-building programme will be further developed in 2018, with a number of courses to be conducted at the regional level.

50. There is tentative agreement for a regional training workshop on the use of satellite imagery data for crop and related statistics in the Asia and Pacific region in the second quarter of 2018. The Queensland University of Technology would take the lead, supported by the Statistics Division. A regional training workshop on the use of mobile phone data for official statistics will very likely take place around the third quarter of 2018. Positium would again take the lead, supported by the Statistics Division. The task team on training, skills and capacity-building will confirm these and further activities in due course.

E. Fifth Global Conference on Big Data for Official Statistics

51. It is proposed that the fifth Global Conference on Big Data for Official Statistics be organized by the Working Group in the fall of 2019 in Africa. In the interim, the Working Group will organize one or more sessions on topics related to its work at the United Nations World Data Forum, to be held in October 2018 in Dubai.

F. Renaming of the Working Group

52. Further discussion will take place on renaming and rebranding of the Working Group and its task teams. Alexandre Barbosa suggested in his keynote speech (see para. 34 above) that the name of the Working Group be changed to bring it more into line with the full scope of activities which the Working Group is undertaking. The new name should reflect the two streams of work of the Working Group, namely the global data collaboratives and the global platform. Furthermore, the task teams could be renamed to reflect their theme or statistical domain rather than the data source, as is currently the case.

VII. Points for discussion

53. **The Commission is invited to express its views on:**

(a) **The progress made by the Working Group on Big Data for Official Statistics and its task teams;**

(b) The outcome of the fourth Global Conference on Big Data for Official Statistics;

(c) The proposal for the progressive development, under the auspices of the Commission, of the Working Group platform as a global collaborative for trusted data and metadata, trusted methods, trusted partners and trusted learning servicing the community of official statistics.

Annex I

Membership of the Global Working Group on Big Data for Official Statistics

Countries

Australia

Bangladesh

Brazil

Cameroon

Canada

China

Colombia

Denmark

Egypt

Germany

Indonesia

Ireland

Italy

Mexico

Morocco

Netherlands

Oman

Pakistan

Philippines

Republic of Korea

Saudi Arabia

Switzerland

United Arab Emirates

United Kingdom of Great Britain and Northern Ireland

United Republic of Tanzania

United States of America

Organizations

African Development Bank

Caribbean Community

Economic Commission for Africa

Economic Commission for Europe

Economic and Social Commission for Asia and the Pacific

Eurostat

Global Pulse

International Monetary Fund

Organization for Economic Cooperation and Development

Statistical Centre for the Cooperation Council for the Arab States of the Gulf

International Telecommunication Union

Statistical Institute for Asia and the Pacific

Statistics Division of the Department of Economic and Social Affairs of the Secretariat of the United Nations

Universal Postal Union

World Bank

Annex II

Text of the Bogota Declaration

Taking into consideration the recommendations contained in the report of the Secretary-General's Independent Expert Advisory Group on a Data Revolution for Sustainable Development, entitled "A world that counts: mobilizing the data revolution for sustainable development", including:

(a) Technology, innovation and analysis to establish a network of data innovation networks for leveraging and sharing data and data research;

(b) Capacity-building and resources related to capacity-building and technology transfer, data literacy and resource mobilization through innovative financing mechanisms in partnership with the private sector;

(c) Governance and leadership related to partnerships and coordination between governments, the private sector, non-governmental organizations, the media and academia to promote good practices and principles in data sharing, open data and data rights,

Recalling the adoption of the Cape Town Global Action Plan for Sustainable Development Data (see E/CN.3/2017/3) and the related outcome documents of the regional conferences on the transformative agenda (see E/CN.3/2017/5) by the Statistical Commission at its forty-eighth session in March 2017,

Highlighting that the Cape Town Global Action Plan for Sustainable Development Data calls upon the global statistical community to take action on the strategic area of modernizing and strengthening the national statistical systems with a focus on modernizing the governance and institutional framework; on applying statistical standards and new data architecture for data sharing, exchange and integration; and on facilitating the use of new technology and new data sources in statistical production processes,

The Global Working Group at the fourth Global Conference on Big Data for Official Statistics in Bogota,

Proposes:

- To provide a major thrust for the strategic area of the Cape Town Action Plan on innovation and modernization by advancing global data collaboratives, facilitated by a trusted federated global platform initially for research and development in the discovery, access and use of data, statistical methodology, software applications and capacity-building for the production of statistics and indicators. These partnerships will innovate and help modernize official statistics and their use of new data sources, including Big Data. It will enable data driven transformation in the production of specific statistics or Sustainable Development Goal indicators for better decision making;

- To progressively invest in research and development via task teams of the Global Working Group whose main objective is the innovation of current statistical production processes and the creation of new ways of compiling Sustainable Development Goal indicators; define a framework for the evaluation of the task teams' work centred on the quality of the information produced as well as their cost effectiveness and scale successful projects by transforming them into data products for global consumption;

- To progressively develop the data and technology architecture of the global platform based on the requirements of the work programmes of the task teams

of the Global Working Group, and underpinned by real user demands and business cases;

- That this work programme for trusted data, services and applications is undertaken under the auspices and guidance of the Statistical Commission, in support of the global community of official statistics by putting the national statistical systems of developed and developing countries at its heart;

- That the global platform should build on the best practices of private and public big data initiatives, offer technology infrastructure and a network for data innovation to the official statistical community and address the need for an interconnected and federated network to facilitate:

  (a) The exchange of ideas and methods for processing, analysing and visualizing big data among official statisticians, data scientists and domain experts from the public and private sectors with a focus on research and development building towards modernized statistical production;

  (b) The sharing and exchange of trusted metadata, methods, services and applications for continuous development and reuse, including sharing trusted data where useful and legally possible. The sharing of sensitive data between trusted partners will be agreed bilaterally, while widely applicable open data sources can be generally made available;

  (c) The development, jointly among the official statistical community, private sector technology companies and other communities, of a trusted data architecture so different types of data can be shared safely and securely;

- To support capacity building via a library of trusted training materials and a catalogue of trusted guidance material, methods and software applications and via conducting workshops on big data and new analytical techniques;

Underlines:

- That the implementation of the global data collaborative as a federated system will place the community of official statistics at the heart of modern trusted data usage and information technology;

- That it will offer both developed and developing countries opportunities to realize the benefits of multisource data, including big data, administrative data, census data and survey data, to better understand economic, environmental and societal changes without investment in expensive technologies;

- That the global collaborative for trusted data, services and applications will benefit all parties involved via synergies in sharing methods and data, creating a global culture of best practice and capability sharing;

- That the sharing of knowledge and capacity-building in human resources in the discovery, access and use of multisource data is a shared responsibility of national and international statistical community and should be scaled in existing statistical capacity-building programmes;

Recommends:

- Global collaboration, facilitated by the global platform with the potential to accommodate many different types of trusted data, trusted services and trusted applications, which should:

  (a) Make it easy for all nations to gain value by participating in the global network;

(b) Deliver a marketplace and a flexible cloud-based technology infrastructure to allow trusted data, methods, services and applications to be shared as a public good where useful and legally possible;

(c) Develop transparent partnership agreements with private and public sector organizations so that network partners contribute and derive value through a business model which is individually sustainable for all stakeholders and ensures access to trusted data.